

Empir Software Eng (2011) 16:460–486
DOI 10.1007/s10664-011-9157-9

Getting the whole story: an experience report on analyzing data elicited using the war stories procedure

Susan Elliott Sim · Thomas A. Alspaugh

Published online: 22 March 2011

© The Author(s) 2011. This article is published with open access at Springerlink.com

Editors: Carolyn Seaman, Jonathan Sillito, Rafael Prikladnicki, Tore Dybå, and Kari Rönkkö

Abstract When analyzing data elicited using the “war stories” technique, previously introduced by Lutters and Seaman (Inf Softw Technol 49(6):576–587, 2007), we encountered unexpected challenges in applying standard qualitative analysis techniques. After reviewing the literature on stories and storytelling, we realized that a richer analysis would be possible if we accorded more respect to the data’s structure and nature as stories, rather than treating our participants’ utterances simply as textual data. We report on five lessons learned regarding how we can better analyze war stories as stories: 1) war stories tend to be about exceptional situations; 2) war stories tend to be diverse and resistant to being combined into a single grand narrative; 3) the humanities can be a valuable resource for analyzing war stories; 4) war stories are not just text, they are also performances; and 5) war stories are not just data, they are also instructive and evocative.

Keywords Qualitative data analysis · Figuration · Methodical · Amethodical

1 Introduction

The war stories procedure, as introduced by Lutters and Seaman (2007), is a data elicitation technique in which the researcher asks participants to recount a highly memorable occasion when adversity was overcome with great effort. After common ground is established, the participant is given a prompt that begins “Tell me about a

S. E. Sim (✉)

Department of Informatics, University of California, Irvine, CA 92697, USA
e-mail: ses@ics.uci.edu

T. A. Alspaugh

Department of Computer Science, Georgetown University, Washington, DC 20057, USA
e-mail: thomas.alspaugh@cs.georgetown.edu

time when” Lutters and Seaman found this procedure elicited longer responses with more contextual detail, but also that the responses were more difficult to analyze and required more space to present. Their paper focused on war stories as a technique for eliciting data from participants, saying relatively little about how to analyze them, beyond referring readers to standard qualitative analysis techniques.

In this paper, we report on our experience analyzing data gathered using the war stories procedure. We asked 34 participants from academia and industry to share stories about how a novice requirements engineer had been detrimental to a project and how an expert requirements engineer had been beneficial to a project. We applied standard qualitative analysis techniques and obtained results that were bland and uninteresting. This took us aback because the raw data had been fascinating and engaging. We decided to reconsider our analysis approach.

A key insight came when we looked at the war stories themselves and how the analysis process was transforming them. Stories have an established structure, and storytelling is a mode of interaction that arguably has always been part of human culture. The standard qualitative analysis process fractures stories into data points, severing structures and relations that made the original story compelling.

For alternative analysis processes we turned to the humanities, an area of scholarly research with extensive experience in drawing larger themes from diverse narratives. In particular, we drew on the technique of *figuration* (Haraway 1997), which collects the rhetoric surrounding a topic to identify a role or frame in which the rhetoric has presented that topic. We used the war stories from our participants to construct a *figure* in this sense of requirements engineering, and this figure led us to insights into the nature of the requirements endeavor. We also employed the *methodical/amethodical* dichotomy presented by Truex et al. (2000). In the context of requirements engineering, the “methodical” is the privileged viewpoint emphasized in texts that describe software development, while the “amethodical” is the marginalized viewpoint that is played down or excluded in these same texts. The methodical viewpoint presents software development as an activity that can be controlled, managed, and regularized. In contrast, the amethodical viewpoint presents software development as an activity that is negotiated, creative, and dynamic. Both facets are present simultaneously in the nature and practice of requirements engineering, but one or the other facet is emphasized by different kinds of presentation. The war stories told by our participants primarily figure the amethodical, while much of requirements engineering research figures the methodical.

From this experience, we present five lessons learned about analyzing data elicited using the war stories procedure. All of these lessons hinge on key insights into the nature and structures of stories.

1. War stories tend to be about exceptional situations, and this lets researchers access data that would not otherwise be available.
2. War stories tend to be diverse and resistant to being combined into a single grand narrative.
3. The humanities can be a valuable source of theories, critiques, and concepts that can be used to analyze war stories.
4. War stories are not just text, they are also performances. Participants carefully select and self-edit the stories that they tell to emphasize points that are relevant to the situation or help them to achieve a goal. We omitted the performative aspect of stories in our study, but will be more aware of it in the future.

5. War stories are not just data, they are also instructive and historical. More so than conventional empirical data, war stories have value beyond the scientific study that elicited them. They can be used to instruct students and to document the history of software engineering. Consequently, consideration should be given to platforms for sharing and archiving of stories.

The remainder of this paper is organized as follows. In Section 2, we present related work on stories and narrative. We describe our study of requirements engineers and the results of our initial analysis in Sections 3 and 4 respectively. Armed with the additional perspectives on stories given in Section 5, we performed a subsequent analysis as described in Section 6. Our lessons learned are given in Section 7, and we conclude the paper in Section 8.

2 Background

A number of disciplines ranging from literature to anthropology and from psychology to robotics have reflected on the nature and structures of stories. In this section, we review selected literature from these fields, as a foundation for understanding war stories not only as empirical data, but also simply as stories. We begin with an examination the role of stories in our culture. Then, we focus on war stories as a genre of stories and as a data collection technique.

2.1 Stories and Narrative

“Story” can be defined many ways but typically involves the connection of subjects and actions in sequence. The narrative mode of relating information is characterized by a *personal viewpoint* (the narrator’s voice), the introduction of *focal actors*, and a *sequence of events* (Pentland 1999). All stories order events; not necessarily chronologically, but also for impact or to explain cause and effect. Stories are especially good for helping the hearer re-experience parts of the teller’s experience, thereby allowing the hearer to re-construct the situatedness of the teller (Schank and Abelson 1995).

Stories have been studied extensively by many different disciplines. Not surprisingly, they are a central concern in literary theory and folklore (Propp 1968). Narratives are rooted in literary theory going back to Aristotle’s investigations of rhetoric and poetics (Aristotle 335 B.C.). In addition, they have been applied in a wide range of fields including organizational theory (Brown and Duguid 2002; Czarniawska-Joerges 1997), psychology (Bruner 2003; Schank and Abelson 1995), and science and technology studies (Haraway 1997). They have even been considered as the human basis of reality and sociality (Bruner 2003; Dautenhahn 2003).

Stories represent an important way of organizing our knowledge about the world. According to theorists, only narrative lets us draw connections and relate how facts or events come together (Bruner 2003). Routine occurrences fit into stories; events and facts that do not fit into the stories are exceptions, to which we should pay attention. The “Narrative Intelligence Hypothesis” goes so far as to say that primate intelligence evolved because we belong to large individualized societies, in which members have formed a social group, but still interact with each other as individuals.

In such societies members need to communicate social dynamics, especially in third-party relationships (Dautenhahn 2003). In contrast, an ant colony is *not* an individualized society; although there is a hierarchy, the ants within each class (e.g. worker or drone) are interchangeable with each other. But in primate groups, members relate to one another individually; a young gorilla's interactions with a bullying, short-tempered playmate differ from those with one who is more placid, easy-going. In order to manage one's relationship with another group member, one may consult a third group member for gossip, support, or protection. Stories are particularly well-suited for encoding and re-constructing socially relevant and meaningful information. Dautenhahn argues that stories are the primary driver for primate brains and primate intelligence (Dautenhahn 2003).

A *script*, as defined by Schank and Abelson (1977), is “a set of expectations about what will happen next in a well understood situation.” For example, a typical interaction between a waitress or waiter in a restaurant and a customer placing an order follows such a script; when each participant speaks, he/she almost sounds as if following lines in a play. Scripts are best suited to understanding stereotypical situations. A person entering a restaurant already knows how to behave in that script, without having to figure out the routine anew. Bruner (2003) relates such phenomena to the idea of *canonicity*. A narrative draws upon a canon of scripts, interactions that are well understood, as its background. But a narrative must break with the canon in some way in order to be “story-like” and worth telling. Scripts are the “necessary background,” but do not define what makes a story. What makes a sequence of events story-like is the inclusion of the unusual, the breach of expectations, or the exception to the rule, something that stands out from the background or creates a new foreground-background relationship.

2.2 War Stories

There are different genres of narrative, such as romance, comedy, horror, and crime fiction. A *war story* is a narrative that presents a non-routine and difficult event, for the purpose of explicating a more general piece of knowledge. A number of previous studies in management and software engineering have used war stories as primary data or the main unit of analysis. McCall et al. (1988) used war stories to study the kinds of knowledge that successful executives learned on the job. Orr (1996) found that war stories were the primary vehicle for sharing knowledge among photocopier repair technicians. Field manuals were of little assistance for diagnosing a photocopier problem, so technicians came to rely on war stories from their peers told as they met over coffee. Learning to state a problem—and its solution—as a war story became a mark of group membership.

Eisenstadt (1997) surveyed USENET newsgroup readers for war stories about their most difficult computer bugs. Sim et al. (1998) used a similar research design and surveyed a similar group of readers for war stories of how they typically and atypically searched for search code. A decade later, Umarji et al. (2008) surveyed participants on discussion boards for war stories of how they searched the web for source code.

War stories as a data collection technique were introduced by Lutters and Seaman (2007). Their article is foundational to our work and we use their technique in our study of expert requirements engineers. The technique consists of an interview with

three phases: *warm-up*, *storytelling*, and *reaction*. In the warm-up phase, a researcher establishes common ground for the topic of the interview, for example, by discussing objectives and definitions. In the storytelling phase, all of the researcher's prompts begin with "Could you tell me about a time when" In the reaction phase, the researcher asks follow-up questions to obtain more detail or to further unpack the story. Lutters and Seaman did not prescribe any particular analysis method, but suggested standard qualitative analysis methods (Miles and Huberman 1994) and the Grounded Theory approach (Strauss and Corbin 1990).

War stories can be viewed as a kind of critical incident technique (CIT). The CIT has been used extensively in management (Gremmler 2004), psychology (Flanagan 1954), and knowledge management (Hettlage and Steinlin 2006), to study human performance, for example, leadership, command ability, and customer service. An incident is "any observable human activity that is sufficiently complete in itself to permit inferences and predictions to be made about the person performing the act" (Flanagan 1954). Criticality has been taken to mean either analyzability or importance. Flanagan defines critical as "...the purpose or intent of the act seems fairly clear to the observer and where its consequences are sufficiently definite to leave little doubt concerning its effects." Roos considers criticality to be defined by the importance of the event and the importance in the memory of the participant recalling the event.

The initial formulation of the CIT in 1954 by Flanagan (1954) had a high degree of specificity with the aim of ensuring repeatability of studies. In contrast, more recent descriptions of the technique emphasize its ability to elicit stories from participants that would not otherwise be available (Hettlage and Steinlin 2006). Despite this evolution, all sources agree that the CIT is suitable for obtaining rich, situated descriptions of both positive and negative exceptions. The CIT is able to gather data not available through other techniques, such as surveys. Unlike other data collection techniques, the effectiveness of CIT is affected by how well the participant can turn a particular sequence of events into a story.

3 Study Design

Our overarching research question was "What is the nature of requirements engineering expertise?" We were interested in the skills and knowledge practitioners typically utilize in carrying out their daily work. In order to answer it, we set up a qualitative study in which we obtained data through interviews that were recorded and transcribed. We initially analyzed the data using standard qualitative analysis methods (Lofland and Lofland 1994; Miles and Huberman 1994) and a Grounded Theory approach (Strauss and Corbin 1990; Corbin and Strauss 2007), organized into the following three standard phases:

1. *Open coding*—The data (interview transcripts, in our case) is *chunked* into segments, each of which is characterized by a single theme or concept, and each segment is assigned a *code* (thematic or conceptual label) that characterizes it. The segments can range from as short as a few words to as long as a complete sentence or several sentences in sequence. The codes arise from what is present in the data.

The segments are compared to identify similarities and differences, and to ensure that each code labels only comparable segments and comparable segments are labeled with the same code. Open coding is iterated repeatedly, re-chunking if necessary, until everything in the data is coded consistently with codes capturing what is important in the data. Throughout this stage, conceptual relations among the segments are uncovered and are expressed through conceptual relationships established among the codes.

2. *Axial coding*—Similar codes are grouped into *categories* and similar categories into higher level categories. The categories are tested against the segments they contain to ensure that each category appropriately represents all segments with codes in that category.
3. *Selective coding*—Candidate *core categories* are identified that represent the main insights, themes, trends, or findings supported by the data. Candidates are evaluated by testing how well each one is supported by the segments and their relationships. An appropriate core category represents at a high level the same conceptual relationships that arose from the codes and categories. The core category then both summarizes and explains the data, and is supported by the data through the framework of codes and categories.

The goal of this standard qualitative analysis approach, and the motivation for the substantial effort required to perform it, is to systematically derive a theme, insights, trends, and relationships at a high level that are well supported by the raw data and the conceptual relationships it expresses.

3.1 Participants

A total of 34 requirements engineers agreed to participate in the study. Participants came from the following three groups:

- 14 attendees at the 2006 International Requirements Engineering Conference (RE'06),
- 15 practitioners at Intuit, Inc. in San Diego, and
- 5 practitioners from elsewhere in Southern California.

As a starting point for distinguishing expert requirements engineers, we used Simon's definition of an expert as someone who has put in more than 10,000 hours of deliberate practice into an activity (Ericsson et al. 1993). This figure translates into roughly 5–10 years of work experience.

Flyers were distributed to all attendees of RE'06, and we interviewed everyone who volunteered to participate. These participants came from several different countries and all were highly experienced in requirements engineering (see Table 1). However, we were concerned that this group was too experienced overall to be representative of requirements engineers as a whole.

In order to obtain a broader experience range, we sought additional requirements engineers from a company in industry, with the goal of finding some novices by Simon's criterion, i.e., three years or so or less of work experience. However, none of this next group of 15 practitioners turned out to be novices, and on average were nearly as experienced as the RE'09 participants. We sought out a final group of five

participants from other companies in order to learn whether this high experience distribution was general throughout industry. The 34 participants are characterized in Table 1.

We found that all three groups of participants were comparable in terms of average years of industrial and RE experience. The RE'06 group averaged 21.4 years in industry and 14.7 in requirements engineering, while the Intuit participants averaged 19.7 and 13.0 years, and the participants from elsewhere in industry averaged 21.6 and 12.2 years. The participants displayed a wide range of educational backgrounds and have held a variety of positions, in both industry and academia. Therefore we argue that the analysis results are not biased toward a particular context or background for requirements engineering.

Although in academia the field is typically referred to as Requirements Engineering and its practitioners as requirements engineers, we (and others with whom we discussed the results) were surprised to learn that *none* of the participants had a job title of “Requirements Engineer”, not even those whose work focused primarily on requirements. The reported titles were Business Architect, Business Analyst, RE Change Agent, Software Architect, Manager, Software Engineer, and Consultant, even for job responsibilities that were clearly in the domain of Requirements Engineering.

3.2 Interview Format

Each interview lasted approximately thirty minutes, and comprised open-ended questions from a script, then follow-on questions for further exploration of issues raised by that participant:

1. What do you think a novice requirements engineer should be able to do?
2. What do you think an expert requirements engineer should be able to do?
3. Please rate your level of expertise.
4. Can you compare what you do now to what you did when you first started out as a requirements engineer?
5. (a) *(If interviewee is an expert requirements engineer)* What advice would you give someone on how to become a better requirements engineer?
(b) *(If interviewee is a novice requirements engineer)* What do you think you would need to learn to become a better requirements engineer?
6. Tell me about a time when involving an expert requirements engineer in a project was advantageous.
7. Tell me about a time when involving a novice requirements engineer in a project was detrimental.
8. Is there anything else you would like to share? Is there a question that you think I should have asked?

Questions 4 and 5 were based on questions used by Campbell et al. (1992) in their study of how programmers proceed through an intellectual developmental sequence as they acquire expertise. Questions 6 and 7 employed the war stories technique (Lutters and Seaman 2007) to probe a phenomenon in context. The interviews were audio-recorded and transcribed, and the transcripts were analyzed.

Table 1 Overview of study participants

Code Name	Occupation (past and present)		Education	Years in Industry	Years in RE	RE-related Publications
	Industry	Academia				
RE '06						
Iona	Consultant	—	Math related MS	42	20+	2
Norah	Consultant	—	Math related PhD	40	15	6
Mark	Software System Engineer	—	Aerospace related PhD	35	20	2
Derek	Software Engineer, Consultant	—	CS/IS related MS, Math related MA	30	20	—
Jacob	Consultant, Systems Engineer	—	CS/IS related PhD	25	10	3
Raymond	Software Engineer, Consultant, Manager	—	CS/IS related MS, MBA	23	7	—
Irwin	Software/Systems Engineer, Consultant	—	CS/IS related MS	20+	10+	9
Edwin	RE Change Agent	—	Math related MS	19	15	3
Anthony	Software Engineer, Consultant, Manager	Lecturer, Postdoctoral Researcher, Professor	CS/IS related PhD	17	29	2
Aaron	—	Professor	CS/IS and Business related PhD	15	10	3
Morgan	Software Engineer, Consultant, Manager	Postdoctoral Researcher	CS/IS related PhD	10	7	—
Muriel	Consultant, Project Manager	—	Marketing/Business related BS	9	1	—
Craig	—	Professor	Psychology related PhD	8	20+	20
Stephen	Consultant, Manager	Professor	CS/IS related PhD	7	22	8
Intuit						
Erica	Manager	—	Math related MS	30+	10	—
James	Business Architect	—	CS/IS related Certificate Program	26	23	—
Malcolm	Software Architect	—	CS/IS related BS	25	20	—
Scott	Business Analyst	—	CS/IS related BS	25	15	—
Dan	Software Architect	—	CS/IS related BS	21	21	—
Myron	Business Architect	—	Telecommunications related MS	21	15	—
Howard	Software Engineer, Manager	—	CS/IS related BS	20+	5+	—
Anita	Manager	—	Economics and Management Science related BS	20	20	—
Bob	Software Engineer	—	CS/IS related BS	20	12	—
Bruce	Manager	—	CS/IS related BS	19	19	—
Joann	Manager	—	CS/IS related MS	18	10	—
Tracy	Software Process Improvement Manager	—	Cybernetics related BS	18	6	—
Audrey	Software Engineer, Manager	—	CS/IS related BS	13	7	—
Jane	Software Designer	—	Industrial Engineering related BS	10	10	—
Carol	Software Engineer	—	EE related MS	10	2	—
Elsewhere in Industry						
Mike	VP of Product Management	—	EE related BS	17	4	—
Kevin	Director of Product Management	—	EE related BS	20	10	—
Ray	Senior Globalization QA Manager	—	CS/IS MS	18	4	—
Mark	Product Manager	—	CS/IS MS	25	25	—
Roger	Director of QA	—	CS/IS BS	28	18	—

4 Analysis Using Standard Qualitative Techniques

In this section, we discuss our initial analysis using the standard approach described in Section 3, and summarize its results.

Our initial analysis identified five traits that were commonly found in requirements engineers.

- *Diplomacy*. This trait is the ability to understand others and make oneself understood, displaying tact, negotiating with care and sensitivity, and the art of saying the right thing at the right time.
- *Approach to Problem Solving*. Expert requirements engineers invest time in understanding the problem, and view this understanding as the fundamental goal of their job. They know from experience that when a problem becomes well defined, good solutions follow easily.
- *Ability to Synthesize Knowledge*. Experts are better able to integrate requirements, knowledge, and teams, even when the system spans multiple domains.
- *Response to Uncertainty*. When the unexpected occurs, experts will adjust quickly and choose a new approach from a diverse toolkit, while novices prefer to stick to a plan, even a failing one.
- *Identifying Stakeholders*. An expert would know when and how to look for a stakeholder with critical information.

In the interests of brevity, we focus on one trait here. We selected Diplomacy because its descriptions in the war stories were diverse and nuanced. The [Appendix](#) contains representative war stories on diplomacy excerpted from four participants' interviews. Each excerpt contains a complete story or piece of advice. They have not been abridged or otherwise edited; these war stories are raw data.

The first thing to notice about these war stories is their length. They are too long to be quoted in full in a conference paper. Furthermore, it is the researcher's responsibility to interpret and analyze qualitative data, to draw larger lessons from multiple data points, so that a contribution can be made to the literature. For both these reasons, we initially analyzed them through a standard process of open-coding, axial coding, and selective coding to identify themes supported by the data.

4.1 Coding

When analyzing qualitative data, one uses a process of constant comparison, in which one is continuously looking for patterns of similarities and differences within and between units of data (interview transcripts, in our case). Open coding is the first step of labeling utterances with topics, followed by axial coding in which topics are grouped into categories. Finally, there is selective coding, in which a category or relationship is selected to be the central organizing principle for the findings. All other categories and relationships provide explanations relative to this central theme.

Coding can start at any time. It is not unusual to begin open coding after two or three interviews are completed. However, we started after the first 14 interviews were completed, since all in that group were collected rapidly at a conference. Similarly, we coded the second and third groups of interviews after each group was completed.

The granularity of the codes will depend on the phenomema under investigation. In this study, we coded at a relatively large scale, because we were interested in traits

Table 2 Open codes for sample diplomacy war stories

Participant	Open codes
Scott	Communication, customer, conflict, relationship, direct
Irwin	Communication, customer, conflict, relationship, not listening
Carol	Communication, customer, expectations, listening
Anthony	Customer, relationship, listening, sensitivity

and characteristics. In a later analysis on the same data, we used finer granularity codes because we were conducting a close reading of the words and concepts that requirements engineers used to describe communication with customers. Table 2 summarizes the codes that we generated. We created the codes inductively as we went along; We did not place an emphasis on having a consistent set of codes that were used repeatedly. If we came up with a code later in the process and realized that it applied to some data that had already been coded, we would go back and re-code.

After we had completed open coding, we proceeded with axial coding, relating the open codes to one another and organizing them into categories. We started with more than a dozen conceptual categories. It was not until we began selective coding and came up with a theme that we arrived at the final set. As with open coding, this was an iterative process of reducing categories and selecting an organizing theme. The final theme was the traits of expert requirements engineers, and the five traits correspond to our final conceptual categories.

As mentioned earlier, we grouped the war stories presented in the Appendix into a category called Diplomacy. It should be noted that this is a conceptual category that we, as analysts, created; it was not mentioned as such by any of the participants, and was identified by comparing war stories across all the participants' data. Also, not every war story in this category had the same codes, and there was no set of codes that were required for inclusion in this category. Finally, a war story could be placed into multiple categories.

4.2 Presentation

Based on our coding, we could identify trends in the data. The following two paragraphs illustrate how we might present such results in a technical paper:

The trait most frequently listed and most strongly emphasized was diplomacy, characterized as tactfulness, skill in negotiation, the ability to understand others and make oneself understood, and the art of saying the right thing at the right time.

Irwin, who has more than 20 years experience in the industry, colorfully stated “A novice can easily cause the blowfish to swell.” He went on to describe one of his early experiences when he was starting out as a requirements engineer. While the customer was explaining a particular point about the technical domain, Irwin interrupted to say “Ahh! So that means [this characterization] as opposed to [that characterization]?” The customer strongly disagreed with Irwin’s characterization, which Irwin felt was a self-evident inference from what the customer had stated. The exchange led to a heated discussion, and thereafter the customer avoided Irwin, hindering progress with the requirements task and reducing the latter’s effectiveness on the job.

But we were haunted by the vividness and compelling interest of the raw data, characteristics which had not survived the qualitative analysis process.

4.3 Categories Versus Stories

We had followed a well-established qualitative analysis procedure and arrived at results that were well supported by the interview data, but were troubled by our realization that these results were uninteresting in comparison with the war stories from which they had been derived. Irwin's full war story in the [Appendix](#) is more compelling than our recounting, even though we include quotes from it such as "A Novice can easily cause the blowfish to swell."

The standard qualitative analysis techniques had produced bland results that were completely at odds with our subjective experience of reading and listening to the war stories. The five traits that were identified (diplomacy, approach to problem solving, ability to synthesize knowledge, response to uncertainty, stakeholder identification, and self-monitoring) were not especially novel and would hardly be considered a substantial contribution to the field.

One could argue that the shortcoming was in our application of the techniques, rather than in the techniques themselves. For example, we could have deepened our analysis to find whether there was any relationship between the participants' backgrounds and experiences and the specific opinions they expressed.

Or one could argue that we were not presenting enough of the participants' evocative statements. Lutters and Seaman (2007) also noted that presenting results from analyzing war stories entailed additional considerations. They wrote "... a study based on war stories does appear to take up more publication real estate in its write-up, especially given that only a tiny subset of the findings [are] presented [there]." In any case, we repeatedly found that while the stories we elicited were rich with insights, the analysis approaches we adopted had failed to preserve them. We began to cast about for an alternative approach.

We realized that the fundamental problem was that the standard qualitative analysis approach of chunking and coding, although it preserved the participant's wording and the conceptual relationships among and across participants' utterances, was discarding the narrative quality of the war stories that was their most compelling feature and from which their richest insights arose. We needed to analyze them in a manner that accorded more respect to the war stories as stories. The participants and their stories were telling us something important, but we needed a different set of analytic tools to hear it. We needed techniques that allowed us to analyze the war stories, but preserved more of their "story"-ness.

5 Perspectives on Stories

In order to capture the most interesting elements of war stories, we needed to find different ways of looking at them. We found two useful dichotomies, text versus performance and modern versus postmodern. The stories that have been elicited can be analyzed as texts, as performances, as grand narratives, or as petit récit. These differences will be discussed here.

5.1 Stories as Text and Stories as Performance

When considering stories merely as text, the focus is on the story as a decontextualized set of utterances (i.e. that which is said by the teller of the story) that convey meaning through their content. The story's meaning is separated from the context of its telling. The story as text is considered meaningful in that it reveals "social facts" or other information about the event or topic being narrated. This is especially the case when researchers ask informants for stories about their experience. There is a tendency to value these stories for their textual components, to consider the "audible story" as a way to access the story behind the story, the real story, such as the events that took place or the information being conveyed. The utterances of the storyteller are taken for granted as communicating the story fully intact, or if the storyteller fails to provide details this is seen as a kind of clouding of the real story which the researcher is trying to access (Boje 1991; Holstein and Gubrium 1995).

Stories tend to be treated as text when they are elicited during qualitative interviews. Utterances are transcribed and coded with an emphasis on the content of the story and the information it conveys or the events that are described. Treating stories as text fits with positivist, evaluative, or normative research interests such as evaluating stories, creating typologies of stories, or discovering and testing individual abilities to recall stories (Boje 1991; Holstein and Gubrium 1995). Such approaches enable researchers to figure out what makes a good story and to abstract out information, story structure, plots, archetypal characters and genres.

When considering stories as performance, focus is placed on the social action of storytelling. From this perspective, stories are highly dependent on the context of the storytelling performance, the specific teller, and their relationship to the audience. For example, the teller of a story may abbreviate the story, or make strategic omissions from it based on the audience (Boje 1991). Listeners are co-producers of stories in that they may interject or "fill in the blanks." Teller and listener will use cues and gestures and pauses to make the story meaningful and to respond to the experience of the story as it unfolds.

Stories are often treated as performances in research approaches such as ethnography and organizational behavioral studies. In such research the storytelling episodes are recorded with the fullest possible contextual information such as gestures and environmental artifacts. Stories are then transcribed with notations for pauses and audience reactions. Analysis focuses on who is telling the story to whom, the messages that might be conveyed by not only the story's content but also the way it is told, perhaps trying to understand the intentions of the teller in telling the story or what it achieves. For example, a story could be told to pass on a lesson, to emphasize common history, or to bully someone into compliance.

Research taking the performative perspective must also assume that stories are influenced by the dynamic between the research analyst and informant. Stories told to the researcher may be shaped for that particular performance (Boje 1991; Holstein and Gubrium 1995). The teller may add details or omit information as a way to demarcate their expertise or distinction. Stories told for others in front of the researcher may not be fully understood by the researcher without extended contact and time spent in situ (Boje 1991; Holstein and Gubrium 1995).

However, both perspectives are necessary to see the kinds of interplay that arise between stories as text and performance. Many storytelling performances include the

invoking of textual understandings of a story, such as in asides like “you know the story” or “so the story goes” (Boje 1991). These are instances when people refer to a “story behind the story,” a decontextualized version, which is then recalled and retold in specific performances. Abstracting and creating the idea of an objective story as text is something that is a part of storytelling performance. In other words, everyone pulls stories out of context and creates stories as text, not just research analysts.

5.2 Modern Versus Postmodern Stories

From the point of view of stories, modernism is characterized by a set of grand narratives such as narratives of progress, justice, and rationality. These modernist narratives are distinguished from narratives of earlier periods during which notions of the subject which drew upon narratives of divine origin, inspiration, and destiny. Modernism, however, replaces these with a new set of grand narratives founded upon rationality and methodical and scientific approaches to knowledge production. These narratives are grand in that they are overarching meta-narratives explaining the human relationship to origin, knowledge, theory, and meaning-making in the world. The Enlightenment narrative, for example, is a grand narrative of origin and telos that claims that human civilization progresses toward a more enlightened state and that introduction of reason and the scientific method progresses human understanding in a linear and cumulative fashion. Rationality was theorized during this period as the highest form of mental functioning. The sciences that we know today developed during the modern period, i.e. the sciences of nature, mind, body, and culture. From a modernist perspective, these are understood to produce universal truths about their subjects.

The postmodern perspective, on the other hand, is marked by an increased skepticism towards and questioning of grand narratives of any kind. Rather than replacing one set of grand narratives for another as was done in the modernist period (for example, the Enlightenment grand narrative of “reason and the historical progression” replaced the older one of “the soul and the divine”), postmodernism rejects the notion of the grand narrative as a way to access truth about the world, about culture, or about the human mind and spirit. A grand narrative is thus understood from this perspective as a story that a culture or society tells itself about its practices and beliefs. Postmodernism is skeptical, for example, of the scientific narrative that “scientific truths are universal and eternal.” Work such as *The Structure of Scientific Revolutions* by Kuhn (1970) identifies and critiques these narratives of science and reveals that scientific knowledge does not progress in a linear fashion. Instead, small narratives within specific communities emerge over time in competition with other narratives and only become standard texts through a complex process of negotiation and consensus-building in which social as well as technical values shape what counts as knowledge.

The postmodernist perspective on narratives is that all stories are multiple, fragmentary, and diverse. The focus is on the “petit récit” or small narratives instead of the grand (Lyotard 1993). While grand modernist stories emerge they are only performances that accomplish specific goals for specific individuals or groups. The scientific method is one of many minor accomplishments of scientists who must not only produce knowledge but produce knowledge as if it adhered to the scientific

Modernism	Postmodernism
<p>Rationality is the highest form of mental functioning</p> <p>Modern science produces universal truths ...independent of context and status of scientist who produced them</p> <p>Rationality will always lead to progress and perfection</p> <p>Reason is the ultimate judge of what is right (true, legal, ethical,...)</p> <p>Language must be rational ...exists only to represent the real world ...the meaning cannot depend on the audience</p>	<p>Questioning the grand narrative (a story that a culture/society tells about its practices and beliefs)</p> <p>Looks instead for mini-narratives (<i>petit récit</i>) that explain small practices, local events, situated, contingent behavior</p> <p>Doesn't make claims about universality, truth, or stability</p> <p>Language is intersubjective ...examines what a text does not say, what it represses (literary deconstructionism) ...look for internal arbitrary hierarchies and dichotomies</p>

Fig. 1 Comparison of modernism and postmodernism (adapted from Steve Easterbrook)

method. Stories that explain small practices, local events, situated and contingent behavior gain favor in postmodernism and are circulated and exchanged. Postmodern research in the humanities and social sciences focuses on critiques of modernist narratives through descriptions of situated action and local knowledge production and the accomplishment of work in varied contexts without making claims about universal truth or stability. Differences between modernism and postmodernism are summarized in Fig. 1.

Literary Deconstruction arose as a scholarly practice in the postmodern period. In deconstruction texts are treated not as containers of information and knowledge but are examined for what they do not say and what they repress—silences and omissions. Deconstruction reveals internal arbitrary hierarchies and dichotomies to understand what they reveal about the performance of the text and what its author was trying to accomplish in relation to a particular readership, often having discover these through close readings of texts in the face of posthumous readings of historical texts. Semiotic theories also shift from structuralist discussions about significations to the relationships between signs and the things they signify—studying questions of how words come to mean what they mean in particular performances and contexts (Geertz 2002; Kay and Kempton 1984).

6 Analysis Informed by Humanities Critiques

As we undertook our second analysis of the war stories, we knew that we wanted to end with more than a pile of facts, that we wanted to arrive at conclusion that would

be a contribution. There were three problems that we had to overcome. In Section 4.3 we discussed the first, the need to avoid making fascinating accounts boring.

Secondly, we had to tread carefully because the central lessons in many of the war stories seemed to go against the dominant paradigms in requirements engineering. Current textbooks and research focus on notations, abstractly-grounded models and techniques, and methods that can be stated, taught, and (presumably) followed. Our respondents gave much greater weight to “people problems.” Methods, models, and notations were mentioned rarely and described as of less importance. These results would be difficult to present and publish in the requirements engineering research community. They would even be difficult to express coherently in the dominant paradigm in that community. We were forced to take a step back and find a different point of view, one from which the results could make sense.

Finally, we made an explicit decision to focus on developing an argument, grounded in the study results, that would convincingly support a conclusion meaningful to the requirements engineering research community. We knew that any such conclusion would have to be supported by some argument, and used potential arguments for candidate conclusions to steer the evolution of both toward better conclusions with stronger arguments supporting them. This approach was influenced by that of Haley et al. (2008) in a different context, in which they guide development of a system and its security requirements by iteratively constructing and evaluating arguments that the system can meet those security requirements.

Figure 2 diagrams an informal argument used in guiding this work. Nodes in the diagram represent waypoints in the argument:

- questions to be answered;
- insights we felt were significant;
- candidate conclusions;
- grounds from our data and analysis;
- potential answers to questions posed; and
- intermediate statements steering the course of the argument.

Edges in the diagram represent paths of reasoning from node to node. We traversed and evolved the diagram repeatedly, seeking out questions that would naturally occur to an audience, answers that addressed them, edges that did not represent convincing steps of reasoning, and conclusions that did not fully express what the preceding argument supported. Where the available conclusions did not convey the full insights and compelling interest of the stories, the difference between what the stories provide and the argument at that time used guided us toward a richer interpretation. As we evolved the argument, this process replaced its conclusions with progressively more interesting ones and its grounds, intermediate stages, and chains of reasoning with progressively more convincing ones.

As discussed above, a breakthrough in our analysis came when we applied some critiques from the humanities. The particular ideas that we drew on were figuration and the identification of privileged versus marginalized texts.

Figuration is the process of collecting up the rhetoric and tropes surrounding a topic or phenomenon into a “figure” as in a role (e.g. figuring in a play) or as a sketch (e.g. a figure drawing) (Haraway 1997). In this sense, figuration is a theory of representation that is performative, situated, and embodied, rather than literal or realist. Haraway developed this analytic technique for clarifying subject-object

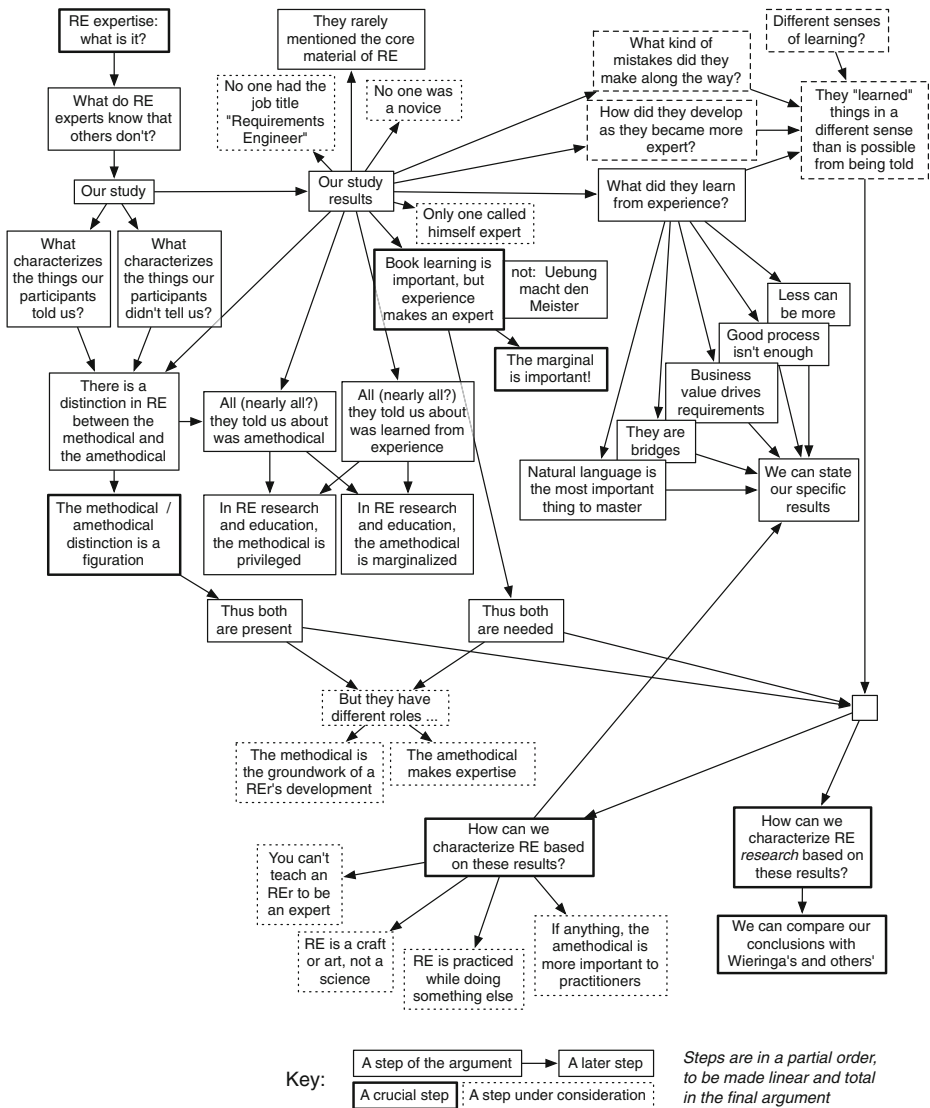


Fig. 2 Diagram of argument: a stage in its evolution

distinctions, particularly in the history of science and technology. She uses figuration because it allows her to hold on to contradictions and heterogeneity in her analyses (Schneider 2005).

The categories of privileged and marginalized emerge from deconstructionism, which is a postmodern analysis technique. By deconstructing a text, one can identify the primary or privileged perspective, that is, the set of beliefs and values that underlie the explicitly stated message. By the same token, the marginalized perspective is the one that is excluded, either implicitly or explicitly, by the privileged

perspective. For every privileged interpretation, there is a corresponding marginalized interpretation, consisting of assumptions and ideals that are deferred as a result of being backgrounded. Truex et al. (2000) described the privileged interpretation in information systems development as *methodical*, that is, a view of the world as ordered, rational, and logical. They describe the corresponding marginalized interpretation as *amethodical*, that is, a view of the world as capricious, random, and socially constructed. In this distinction, “amethodical” does not mean careless or without procedure, but rather beside or outside of method.

Both of these critiques are helpful because they show how competing narratives can be simultaneously true. Figuration allowed us to collect up the different views of requirements engineering, as given by participants in our study, and the competing ones found in the textbooks and scholarly manuscripts. This technique showed us how bits of stories and tropes could be put together into a figure. The categories of privileged and marginalized provided us with an example of how to accommodate contradictory figures or perspectives within a single analysis or argument. This critique showed us how to use what was said to frame what was unsaid. By combining figuration with privileged/marginalized, we were able to arrive at a new way of looking at requirements engineering research and practice, one where conflicting stories could be true simultaneously.

A *methodical figuration of requirements engineering* brings certain tools and techniques into the foreground: those that regularize requirements, use formalisms, and employ logical decompositions are valued and given prominence. Examples are formal methods and model checking approaches that exploit the rationality underlying software systems, and techniques for imposing order such as ontologies and XML. “Soft issues” are pushed to the background and marginalized (Goguen and Linde 1993; Viller and Sommerville 1999).

If we instead adopt an *amethodical figuration*, other aspects of requirements engineering are placed in the foreground. In this perspective, the world is viewed as negotiated, capricious, fragmented, and creative. A different set of tools, notations, and techniques are valued. Examples are contextual inquiry (Beyer and Holtzblatt 1997) and workshops to increase creativity in requirements (Maiden et al. 2007); for tools, approaches such as Chechik et al.’s multi-valued logic work (Chechik et al. 2003) and work by Sabetzadeh and Easterbrook (2006) on merging multiple, sometimes inconsistent models.

Each of these two figurations makes one aspect more prominent. Both aspects are always present, but one is made easier to look at than the other. The figuration approach lets us hold a more inclusive picture of requirements engineering in our mind’s eye.

All five of the themes we identified were either *amethodical* or highlighted the tension between *methodical/amethodical*. This contrasts with the six traits found in the first analysis in Section 4, all of which fit into the standard *methodical* view. A more detailed presentation of this analysis can be found elsewhere (Sim et al. 2008).

1. *Requirements engineers are bridges between worlds*

The worlds to be bridged typically vary from project to project, as do the ways in which they can be bridged. The statement that worlds need to be bridged is itself inherently *amethodical*, since in the *methodical* view there are not multiple worlds and a rational system needs no bridging.

2. *Good communication is key*

Good communication appears to be primarily amethodical, in the sense that every good communication has “a unique and idiographic form” chosen for the situation, the participants, and the matter to be communicated.

3. *Good processes help, when used selectively*

Processes are methodical by definition, but the *selective* use of process where appropriate, described by the study participants, is the result of experience and judgement. The participants also state that using a good process is not sufficient to achieve a good result, and the additional desiderata are amethodical.

4. *With the appropriate abstraction, less is more*

While the choice of an appropriate abstraction is aided by methodical knowledge and techniques, the selection and use of the abstraction relies on the amethodical, and the idea of valuing a less-detailed requirements document contrasts with the methodical view in which a single comprehensive document is preferred.

5. *Business value, not technical elegance, should drive requirements*

Business value is always situated and ad hoc, and varies from situation to situation. In contrast, technical elegance arises from general principles that are rational and universal. Favoring business value is amethodical, while favoring engineering process is methodical.

In each of these cases, we see that a substantial part of the aspect described is amethodical. As noted by Truex et al., the amethodical is a concept that is marginalized in methodical texts. Requirements engineering books and academic courses tend to focus on general principles seen as of lasting value, in other words, the methodical. Combining Truex et al.’s analysis and Haraway’s technique of figuration, we arrive at the idea that when a marginalized concept is brought to the foreground, it is not with the intent of replacing a privileged text, but to introduce a new way of understanding that shifts fluidly between alternative figurations.

7 Lesson Learned

We have distilled five lessons from our experience analyzing war stories. We found that war stories were difficult to analyze. As illustrated by our experiences discussed in this paper, qualitative analysis techniques that are commonly used in software engineering (Miles and Huberman 1994; Seaman 2010) could only get us so far. We had to re-group and reflect on our methods and goals in order to arrive at the result that was finally published.

7.1 War Stories Procedure Elicits Data About Exceptions

One of the most significant strengths of the war stories procedure is its capacity to solicit data from off the beaten track. Because the prompts are open-ended and place few constraints, participants can respond with any story that they feel is worth telling. We have identified two reasons for this capacity of the war stories method.

First, the stories that are solicited are less influenced (for good or ill) by the researcher’s theoretical biases or conceptual framework. In our study as well as the study by Lutters and Seaman, warm-up questions were used to set the stage

and establish a topic for the interview, but few other occasions arose to direct the conversation in one way or another. For comparison, consider the open ended questions that are used in an interview as part of an exploratory study. These questions are designed to solicit information on a particular topic. The interview script ensures that data is collected on the same topic from each participant. No such framing accompanies the war stories prompt; there is no safety net to ensure that everyone provides data for every variable of interest in a structured, regular manner.

Second, good stories are about exceptions, rather than the rule. A story that is about an event that happens regularly without exception is not worth telling. This is not the stuff of movies, novels, or folk tales. Only when something unusual happens does the story progress. As observed by Schank, stories help us to identify and contextualize novel information. In order to know what is unusual or extraordinary, we must first know what is routine or expected. Scripts provide the background, so that breaches, violations, and innovations can be made into stories with rising and falling action.

Generalizability and repeatability of the war stories approach may appear to be a concern, but these criteria belie a realist or positivist philosophical perspective commonly associated with quantitative methods. It would be more appropriate to evaluate the war stories approach using criteria from a constructivist perspective better suited to qualitative methods, in particular, transferability and dependability (Guba and Lincoln 1989; Trochim 2000). Transferability is judged by the person trying to take the results into another setting. Dependability comes from the researcher describing the ever-changing context in sufficient detail and accounting for the effects of these changes on the observations.

To use the war stories procedure effectively, researchers will need to scope the phenomenon of interest appropriately to avoid being overwhelmed by the exceptional cases in the data. Our interest in requirements engineering expertise was comparatively broad and unwieldy, while Lutters and Seaman's interest in documentation use was narrower and more suited to (modernist) standard data analysis techniques. They were interested in successful and unsuccessful attempts to locate and use documentation in the context of software maintenance. These tended to be discrete events. In contrast, requirements engineering is conceptually ambiguous and an activity spanning a period of time. The work of a requirements engineer can vary greatly depending on the requirements engineer's expertise, the scope of the task, and the nature of the customer relationship. It is easier for a participant to tell a story about a specific incident than to summarize a sequence of events over time.

7.2 Fitting Many War Stories into One Big Picture is Difficult

When analyzing data from an empirical study, the goal is usually to arrive at a smaller number of statements that summarizes the observations. In other words, the objective is to find regularities, overarching themes, or even a theory. However, this task can be difficult since the data produced by war stories tends to present exceptional situations or events.

In our initial analysis, we distilled the data using both quantitative and qualitative techniques. We calculated the participants' average age, average number of years

of experience working with requirements, the number of articles they published in top research venues, and their educational background. We identified common themes in the war stories by inductively identifying variables of interest and coding the data. In other words, we followed all the usual steps for analyzing interview data. Unfortunately, this approach seemed to take all the richness out of the stories. An analysis that looks for commonalities by its nature filters out the exceptions that are critical to making stories lively and compelling.

It is difficult to explain how a particular narrative is engaging without sharing the narrative itself. The value of a story is most accessible when the story is intact. Stories are a form of communication that has been with us for a long time. As such, we have learned both how to tell and how to hear stories. When we take apart a narrative without regard to its overall story arc and relationships among the elements, something is lost. This is clearly a case where the whole is more than the sum of the parts.

There is a conflict between the need to build a coherent body of knowledge, and the desire to embrace each war story as a stand-alone lesson. Data analysis typically requires that we build some kind of generalization or larger analytical story that encompasses the many data points we have collected. Data points like the stand-alone war story should be put together through scientific analysis into a larger story that can be added to the knowledge base of the discipline. However the war story, rather than fitting in to this background knowledge, stands out from it and is difficult to integrate. This conflict between the needs of the discipline and the stand-alone war story is similar to that between the modernist and postmodernist perspectives. According to the modernist perspective, singular events can be made to fit into a grand narrative that ties everything together. Science and logical positivism are essentially modernist projects in that they seek to obtain data that can be strung together to form an idea about the single objective world. Empirical observations are merely points of data that do not cohere until they are brought together to form a universally applicable theory about that underlying reality.

In software, the influence of postmodernism is clearly present, but is not always explicitly acknowledged. For instance, there is a separation between software process models and their enactments (Fuggetta 2000). Process models are like a grand narrative of modernism whereas research on software process enactments demonstrates that every software project is unique and stands alone. In requirements, we take care to obtain feedback from diverse stakeholders (Sharp et al. 1999), recognizing the value of many points of view, rather than the objective scientific point of view. When conducting a formal software inspection, we may use a technique that requires us to take on a particular perspective, such as testing (Basili et al. 1996). In other words, our research and methods acknowledge that multiple true stories can exist simultaneously, yet we do not always know what to do with them, and in particular how to share them with the research community.

In software, one common way to bring together many *petits récits* is to identify best practices, which are recommendations or lessons that can be learned from others' experiences. Unfortunately, it is difficult to identify best practices using only exceptional cases, elicited as war stories. Nevertheless, war stories, and more broadly the exceptions, are still meaningful and worthy of analysis, so we must find ways to learn from them. In the next subsection, we will describe how we successfully borrowed critiques from the humanities for this purpose.

7.3 Concepts from the Humanities can be Helpful when Analyzing War Stories

We found that the humanities were a valuable source of critiques and analytic tools such as figuration, post-modernism, and protocol analysis, because much attention has been given in that field to finding ways to analyze stories while still preserving story structure. In the humanities, scholars seek to understand the human condition; critiques and artistic contributions that have more explanatory power are valued more. Critical, Marxist, and feminist critiques are part of a historical succession of analytical approaches that seek to more faithfully describe and explain our diverse experiences as human beings.

In contrast, in the sciences only provable facts are true. Sometimes humanistic or hermeneutic truth is disparaged by scientists as subjective, culturally relative, or even vague, because arguments are not boiled down to a single point and the form of the argument is just as important as the conclusion. But these are precisely the reasons why concepts from the humanities can be helpful when analyzing war stories.

It is not a trivial task for a software engineer to become educated in a spectrum of theories from the humanities. We became aware of these concepts through collaboration with colleagues in the humanities, attending seminars, taking courses, and participating in reading groups. In other words, we became participants in a collaborative interdisciplinary scholarly community. Our own experience led us to realize that the learning curve is steep for those analyzing war stories and we have no easy solution. We report this lesson learned to point out the possibility of borrowing from the humanities when other analytic lenses, such as those from computer science or social science, fall short.

7.4 War Stories are Performances Too

In our analysis, we treated the war stories only as text, which was what we were taught to do in survey research. We tried to avoid leading questions so as not to bias the participants. We knew from textbooks that personal characteristics could be a threat to validity, but accepted this as part of the territory since there was little that we could do about it. There was no way to obtain the data without *somebody* sitting down with the participants and asking questions.

With our recent, deeper understanding of storytelling and narratives, we realized that we underestimated our own role in the data elicitation. We neglected entirely the fact that stories are performances too. People choose which stories to tell, and edit a story, not to manipulate, but to emphasize sequence, and cause and effect. Stories are told for a purpose. In retrospect, we see that the participants were also performing stories for us. At times in the interviews, the participants would address us directly as researchers who were young (or at least youthful in appearance). It is also highly likely that the participants were performing, in the sense of deporting themselves, as expert requirements engineers and were speaking to us in a deliberate manner.

Our personal equations may have led the participants to choose certain kinds of stories to tell, which led to the results that we obtained. Due to the nature of storytelling, the participants likely selected stories that they thought were novel to us or complemented book knowledge. We have no evidence to argue one way or

another, but it is possible that they were trying to encourage research or education in requirements engineering in a certain direction. Part of the postmodernist perspective is that perfect evidence is impossible. Knowing why and how participants chose particular stories to tell and what counterfactual stories they might have performed in other circumstances is done by reading between the lines, not by seeking hard evidence.

None of us can escape our personal characteristics when collecting data, so we need to be aware of how we influence the war stories that are told. This influence is present generally when interacting with participants (Holstein and Gubrium 1995), but we suspect the war stories elicitation technique is particularly sensitive, because storytelling is performative. We can pay attention to how the interviewee is addressing and responding to us as researchers. From a modernist perspective, it is desirable to mitigate this influence, possibly by having two researchers who are demographically or philosophically different from each other conduct the interview together. However, from a postmodernist perspective it is not necessary to remove the bias, because it is not a contaminant, but a characteristic of the data. Ultimately, we need to recognize that individuals and interactions are part of how all knowledge is constructed and that there is no underlying true and unadulterated knowledge to access by making methods less biased or less performative.

7.5 War Stories can be More than Data

When working with war stories, we often wished that we could share the entire story or interview. We found the narratives to be very compelling and highly instructive. When we were listening to the recordings or reading the transcripts, we often became caught up in the story and sometimes lost track of our original task. We suspect that Lutters and Seaman (2007) had a similar experience because, as noted above in Section 4.3, they found themselves including longer quotations than usual in their article.

We felt that both the research community and students new to requirements engineering would benefit from hearing the stories. The stories have educational value, because movement forward in the story is closely tied to the lesson at the end, which causes students to become engaged with the material and to attend to the appropriate details. Researchers would also benefit from hearing stories, because they describe aspects of requirements engineering that they do not interact with regularly. While social factors are widely acknowledged, there is a large gap between theoretical knowledge and a concrete example.

Outside of software engineering, efforts are being made to record and share stories. In history and humanistic disciplines, “oral histories,” which are first-person accounts of historical events, are collected and archived. StoryCorps is a non-profit organization that travels across the USA to gather stories from average people as told in conversation with a friend or loved one. The stories are subsequently archived in the American Folklife Center in the Library of Congress. Following this example, two consulting companies teamed up to support the “Agile Corps” project (Agile Corps 2008) to capture stories from people who have worked with, taught, or invented agile software development methods.

These experiences suggest that war stories should be shared in ways that other kinds of research data are not. In order to expand the life and uses of war stories,

these possibilities need to be considered from the outset. Informed consent needs to be obtained from participants prior to data collection, which in turn requires this step to be written into study protocols and approved by institutional review boards. After the stories are gathered, sharing them is not trivial issue. Additional effort is needed to publish, label, and edit the recordings and/or transcripts. Ideally, a carefully designed platform for publishing, sharing, and consuming war stories could be implemented. A lively war stories archive is a non-trivial amount of work and is unlikely to happen without significant community support.

8 Summary and Implications

The war stories procedure is a robust and flexible data elicitation technique. They gather rich, contextualized accounts of memorable events from participants. We used this technique to gather data on requirements engineering expertise from industrial practitioners and academic researchers. This paper reports on our experiences and lessons learned from analyzing this data.

We provided two analyses of the war stories to serve as illustrative examples. The first analysis used qualitative data analysis techniques that are commonly employed in mainstream software engineering. These techniques fractured the stories into facts (which led to the loss of the story structure) and then proceeded inductively (without the perspective provided a theory). The result was an analysis that was bland and lacked enough substance to be a contribution to the literature. The second analysis used humanities analysis techniques that allowed us to preserve more of the story structure. In addition, we used the evolution of the argument we proposed to make in presenting the analysis results to guide us in taking full advantage of the rich, evocative war stories. These approaches, along with the addition of concepts such as methodical/amethodical, yielded an analysis that gave us new insights into requirements engineering research and practice.

Based on our experience and a review of the literature on stories, storytelling, and narratives, we identified five lessons learned.

1. War stories tend to describe exceptional situations.
2. War stories tend to be diverse and resistant to being combined into a single grand narrative.
3. The humanities can be a valuable source of techniques and critiques that can be used to analyze war stories.
4. War stories are not just text, they are also performances.
5. War stories are not just data, they are also instructive and historical.

Taken together, these lessons suggest that we are only beginning to get the whole story from war stories. Our work contributes to a burgeoning body of knowledge on how to use them effectively. There is little doubt that war stories provide unique access for researchers to software engineering in action. But are many more potential analyses and applications for war stories. The possibility of archiving and sharing war stories as a kind of oral history is intriguing. A new set of data collection and curation practices would be required, but once established, an archive of war stories would have broader significance for both research and education. The research community

should continue to explore and evaluate the use of war stories in empirical software engineering.

Acknowledgements Our first debt of gratitude is to the participants in the study who shared their experiences with us. We spent many delightful hours listening, reading, and working with their war stories.

We would like to give a special thanks Ban Al-Ani for her contributions to conducting the original study of requirements engineering expertise. The study would not have happened without her. Thanks to Medha Umarji who helped to develop the initial survey instruments. We are grateful to Vivian Olivera and Joey Lei who helped conduct and transcribe interviews, and Swaminathan Subramaniam who also helped to transcribe the data. We thank Marisa Cohn for help with background research on narratives.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

Appendix: “Diplomacy” War Stories

Scott

I was in a healthcare company that had under a 1,000 employees all housed in the same building. The requirements engineer was working with the customer directly, a senior manager in this particular instance who was very direct. And you had to understand that direct was the way they operated and you couldn't come in and not expect to have very straight forward and matter of fact conversations. They started working with the requirements and got into an escalation match. They got into this contest that after an hour they certainly had hacked out a set of requirements. However, that session had poisoned the relationship because both sides became competitive and unfortunately that resulted in a phone call to the other side saying that “what is this guy doing? He's causing havoc and blah blah blah.” It was unfortunate cause while a person is direct, there are ways to re-frame questions. Instead of being direct right back, you let them come out with it. This person just didn't know that. What happened was we had to take this requirements engineer and put his on some other project and get another senior person and repair the relationship and actually get to some of the things that the customer was trying to get to but the previous guy didn't have the experience and didn't recognize the situation and got into the wrong kind of dialog. It was BAD.

Irwin

When I was just starting out, I did some interviews with a media company and I think there was one guy who was explaining something about himself. He was explaining something about a part of the technical domain. And I did something fairly chin first like “Ahh so that means this category as opposed to that category.” And he basically violently disagreed with what was more or less self evident from what he had said the sentence before. So it is certainly possible for a novice merely by reflecting back to people or by pursuing back a

reflecting to people slightly vigorously. You can rehash things and make a hash of them. And people can get cross at you asserting something. Whereas, if you just delicately say, “Does that mean that?” Or “Could you explain that?” By elegantly going around things it is possible to avoid getting any kind of stressed out reaction from people. People can be very sensitive. Suzanne says that it is surprisingly easy to upset people.

Carol

Learn the technique of active listening. Make sure your understanding is the same as the customer’s understanding and also set the right expectations. I still think that domain knowledge is important. If you don’t know what you are talking about then chances are that you are not asking the right questions either.

Anthony

I learned the business of dealing with customers through him. I basically shadowed him. I told him that I was shadowing him and he let me shadow him. We went out to, maybe, 15 customers over a period of six months. I let him do all the talking and he let me do the talking when it was appropriate. And I was able to see how *tender* he was. Here’s a PhD in mathematics, who was *tender* with his customers. He *cared* about his customers. And that told me that, oooh, being a PhD doesn’t mean that you have to be a hard ass. You can be a soft person and understand and listen.

References

- Agile Corps (2008) <http://community.serena.com/hives/5651faf9d7>, Accessed 3 Oct 2010
- Aristotle (335 BC) Poetics (trans. SH Butcher). <http://classics.mit.edu/Aristotle/poetics.html>
- Basili VR, Green S, Laitenberger O, Lanubile F, Shull F, Sørumgård S, Zelkowitz MV (1996) The empirical investigation of perspective-based reading. *Empirical Software Engineering* 1(2): 133–164
- Beyer H, Holtzblatt K (1997) Contextual design: defining customer-centered systems. Morgan Kaufmann, San Mateo, CA
- Boje DM (1991) The storytelling organization: a study of story performance in an office-supply firm. *Adm Sci Q* 36(1):106–126
- Brown JS, Duguid P (2002) The social life of information. Harvard Business School Press
- Bruner J (2003) The narrative construction of reality. In: Mateas M, Sengers P (eds) Narrative intelligence, chapter 3, pp 63–90. John Benjamins Publishing
- Campbell RL, Brown NR, DiBello L (1992) The programmer’s burden: developing expertise in programming. In: The psychology of expertise: cognitive research and empirical AI. Springer-Verlag, pp 269–294
- Chechik M, Devereux B, Easterbrook S, Gurfinkel A (2003) Multi-valued symbolic model-checking. *ACM Trans Softw Eng Methodol* 12(4):371–408
- Corbin JM, Strauss AC (2007) Basics of qualitative research: techniques and procedures for developing grounded theory. Sage Publications
- Czarniawska-Joerges B (1997) Narrating the organization: dramas of institutional identity. University of Chicago Press
- Dautenhahn K (2003) Stories of lemurs and robots: the social origin of story-telling. In: Mateas M, Sengers P (eds) Narrative intelligence, chapter 4. John Benjamins Publishing, pp 63–90
- Eisenstadt M (1997) My hairiest bug war stories. *Commun ACM* 40(4):30–37

- Ericsson KA, Krampe RTh, Tesch-Römer C (1993) The role of deliberate practice in the acquisition of expert performance. *Psychol Rev* 100(3):363–406
- Flanagan JC (1954) The critical incident technique. *Psychol Bull* 51(4):327–358
- Fuggetta A (2000) Software process: a roadmap. In: *Proceedings of the conference on the future of software engineering*. ACM, pp 25–34
- Geertz C (2002) Thick description: toward an interpretive theory of culture. In: *Culture: critical concepts in sociology*, pp 173–196. Routledge
- Goguen JA, Linde C (1993) Techniques for requirements elicitation. In: *Proceedings of IEEE international symposium on requirements engineering*, 1993, pp 152–164. doi:[10.1109/ISRE.1993.324822](https://doi.org/10.1109/ISRE.1993.324822)
- Gremler DD (2004) The critical incident technique in service research. *J Serv Res* 7(1):65–89
- Guba EG, Lincoln YS (1989) *Fourth generation evaluation*. Sage Publications
- Haley CB, Laney R, Moffett JD, Nuseibeh B (2008) Security requirements engineering: a framework for representation and analysis. *IEEE Trans Softw Eng* 34(1):133–153
- Haraway DJ (1997) Modest witness@second millenium. *FemaleMan meets OncoMouse: feminism and technoscience*. Routledge
- Hettlage R, Steinlin M (2006) The critical incident technique in knowledge management-related contexts. Technical Report, IngeniousPeoplesKnowledge
- Holstein JA, Gubrium JF (1995) *The active interview*. Sage Publications
- Kay P, Kempton W (1984) What Is the Sapir-Whorf hypothesis? *Am Anthropol* 86(1):65–79
- Kuhn TS (1970) *The structure of scientific revolutions*. University of Chicago Press, Chicago
- Lofland J, Lofland LH (1994) *Analyzing social settings: a guide to qualitative observation and analysis*. Wadsworth
- Lutters WG, Seaman CB (2007) Revealing actual documentation usage in software maintenance through war stories. *Inf Softw Technol* 49(6):576–587
- Lyotard JF (1993) *The Postmodern condition: a report on knowledge*. University of Minnesota Press
- Maiden N, Ncube C, Robertson S (2007) Can requirements be creative? Experiences with an enhanced air space management system. In: *28th international conference on software engineering (ICSE '07)*, pp 632–641
- McCall MW, Lombardo MM, Morrison AM (1988) *The lessons of experience: how successful executives develop on the job*. Simon and Schuster
- Miles MB, Huberman M (1994) *Qualitative data analysis an expanded sourcebook*, 2nd edn. Sage Publications
- Orr JE (1996) *Talking about machines: an ethnography of a modern job*. Cornell University Press
- Pentland BT (1999) Building process theory with narrative: from description to explanation. *Acad Manage Rev* 24(4):711–724
- Propp VIA (1968) *Morphology of the folktale*. University of Texas Press
- Sabetzadeh M, Easterbrook S (2006) View merging in the presence of incompleteness and inconsistency. *Requir Eng* 11(3):174–193
- Schank RC, Abelson RP (1977) *Scripts, plans, goals and understanding: an inquiry into human knowledge structures*. Lawrence Erlbaum Associates Hillsdale, NJ
- Schank RC, Abelson RP (1995) Knowledge and memory: the real story. In: Wyer RS, Schank RC, Abelson RP (eds) *Knowledge and memory: the real story*, chapter 8. Lawrence Erlbaum Associates, pp 1–86
- Schneider J (2005) Donna Haraway: live theory. Continuum International Publishing Group
- Seaman CB (2010) Qualitative methods. In: Shull F, Singer J, Sjöberg DIK (eds) *Guide to advanced empirical software engineering*, chapter 2. Springer
- Sharp H, Finkelstein A, Galal G (1999) Stakeholder identification in the requirements engineering process. In: *Proceedings of the 10th international workshop on database and expert systems applications*, pp 387–391
- Sim SE, Alspaugh TA, Al-Ani B (2008) Marginal notes on amethodical requirements engineering: what experts learned from experience. In: *Proceedings of the 16th international requirements engineering conference*, pp 105–114
- Sim SE, Clarke CLA, Holt RC (1998) Archetypal source code searches: a survey of software developers and maintainers. In: *Proceedings of the 6th international workshop on program comprehension*, pp 180–187
- Strauss AL, Corbin J (1990) *Basics of qualitative research: grounded theory procedures and techniques*. Sage Publications
- Trochim WM (2000) *The research methods knowledge base*, 2nd edn. Atomic Dog Publishing, Cincinnati, OH

- Truex D, Baskerville R, Travis J (2000) Amethodical systems development: the deferred meaning of systems development methods. *Account Manag Inf Technol* 10(1):53–79
- Umarji M, Sim SE, Lopes CV (2008) Archetypal internet-scale source code searching. In: *Open source development, communities and quality*, vol 275/2008. Springer, pp 257–263
- Viller S, Sommerville I (1999) Social analysis in the requirements engineering process: from ethnography to method. In: *Fourth IEEE international symposium on requirements engineering (RE'99)*, pp 6–13



Susan Elliott Sim is an Assistant Professor in the Department of Informatics at University of California, Irvine. Her primary research area is empirical software engineering, with particular attention to program comprehension, source code searching on the web, research methodology, and software process for small companies.



Thomas A. Alspaugh is an Adjunct Professor at Georgetown University in Washington, D.C. His research focuses on requirements methods and tools for getting the right software solution to the right human problem. Prominent themes within his work are the interactions between law and requirements, OSS licenses, privacy, and effective use of scenarios and narrative requirements.